DOCUMENT RESUME

ED 259 026 TM 850 398

AUTHOR Alford, David W.
TITLE IQ Test Controversy: Past, Present, and Future Trends.
PUB DATE [84]
NOTE 24p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Culture Fair Tests; Educational Trends; Elementary Secondary Education; Futures (of Society); Intelligence; Intelligence Differences; *Intelligence Quotient; *Intelligence Tests; Nature Nurture Controversy; Response Style (Tests); Test Bias; *Testing Problems; *Test Use

ABSTRACT

The controversies surrounding the use of intelligence quotient (IQ) tests with children are summarized. This article discusses what intelligence is and how intelligence is measured. It also examines factors which can affect measurement, including examiner training or bias, examinee age, misinterpretation of test scores, and poor tests. The appropriate uses of IQ tests are considered, as well as the abuses. Past, present, and future trends in intelligence testing are described. It is concluded that since trends suggest that more rather than less IQ testing will occur, efforts should be directed at developing better tests and reducing abuses. (GDC)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED259026

# IQ TEST CONTROVERSY

## PAST, PRESENT, AND FUTURE TRENDS

David W. Alford, Ph.D

D'Youville College

Buffalo, New York

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Alford, David W.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 850 398

ABSTRACT

The article reviews the controversy surrounding the use of IQ tests with children. Included is a discussion of what intelligence is and how it is measured, along with an examination of factors which can affect measurement. The appropriate uses and the abuses of IQ are considered as are the past, present, and future trends regarding IQ measurement.

> There is an old Army saying that, "if it moves salute it; if it stands still paint it." Today, some reformers seem to be saying, "if it moves, test it." (Currence, 1984, p.33)

Controversy has surrounded the concept of intelligence measurement almost since the development of the first IQ test by Alfred Binet in 1905. Periodically, articles about IQ tests and measurement appear in both professional journals and the public media. These articles range from hyperbole and condemnation to unqualified support. The use of these tests has been declared illegal in several states if given to minority children for the purpose of placement in a class for mentally retarded children. Organizations such as the National Education Association (NEA) have called for a moratorium on the use of standardized intelligence tests in the schools although many state legislatures are moving in the direction of greater teacher accountability as measured by standardized tests.

In a comprehensive review of the controversy surrounding mental testing which has occurred over the past 50 years, Cronbach (1975) concluded that the criticism of testing often has lacked scientific merit. He said, "all the public hears is endless angry clamor from extremist who see no good in any test or no evil in any test" (p. 1). These extremes are reflected by statements such as that by Kamin (1974) who said, "since its introduction to America, the IQ test has been used more or less consciously as an instrument of oppression against the underprivileged, the poor, the foreign born, and racial minorities" (p. 12). In contrast, for example, is the statement by Sattler (1982) that the development of intelligence tests is one of the most significant contributions made by the field of psychology. The use of IQ tests in the public schools is clearly a controversial topic. Consequently, the purpose of this article is to review the characteristics associated with mental ability tests and, more importantly, to present both sides of the issue -- the use and abuse of IQ tests.

DEFINITION AND CHARACTERISTICS OF INTELLIGENCE

Before continuing a discussion of intelligence testing, it would be appropriate to first consider what intelligence is. There is such a thing as intelligence and almost everyone agrees that intelligence is a good thing to have, yet there is surprisingly little consensus as to what it is. Psychologists have known a lot about the measurement of intelligence but almost nothing about what it is they are measuring (Sternberg, 1979). Intelligence is sometimes mistakenly thought of as a physical attribute much like hair color, or a quantifiable substance like height or weight. This difficulty in defining intelligence is reflected by a symposium sponsored by the Journal of Educational Psychology in 1921 for the specific purpose of developing a definition of intelligence. Thirteen different definitions were developed by the 13 participants although there was some commonality among the answers (Rice, 1979). Even today, psychologists, educators, sociologists, geneticists, neurophysiologists, biochemists, and others still are attempting to develop a workable definition (Weiner & Stewart, 1984).

Binet, the developer of the first IQ test, considered intelligence as a collection of such abilities as judgement, practical sense, initiative, and the ability to adapt to circumstances (Sattler, 1982). Vernon (1969) associated three meanings to the term. One is that intelligence refers to the innate capacity or biology of an individual; another meaning refers to behavior or what individuals do in response to stimuli; and the third refers to results obtained on intelligence tests. Indeed, one somewhat tautological definition is that intelligence is what intelligence tests measure (Sattler, 1982). Perhaps the most operational definition was provided by Wechsler (1958) who stated that, "intelligence is the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with the environment" (p. 7).

BEST COPY

In addition to difficulties in defining intelligence, psychologists also can not reach consensus regarding the extent to which environment affects intelligence. This so-called "nature-nuture" controversy is, in fact, one of the most studied and debated issues in psychology since its implications are enormous. The controversy involves the relative extent to which environment and biology affect intellectual development. The fervor of the debate reached its zenith following the publication of an article by Jensen in the Harvard Educational Review in 1969. Jensen concluded that approximately 80 percent of intelligence results from heritability. Much of the article was based on various studies comparing the intellectual characteristics of identical twins raised both together and apart. Loehlin, Lindzey and Spuhler (1975) summarized a number of these studies and found that as the genetic component between people became more similar, their IQs also became more highly related. For example, the correlation of IQs between unrelated children is .17 whereas the correlation of IQs between identical twins is .88 thus suggesting a genetic factor in intelligence.

Other researchers (Mercer, 1979; Scarr, 1978; Vernon, 1979), however, contend that environment, not biology, has the greatest influence on intellectual development. One classic study often cited is that of Skodak and Skeels (1949) who transferred "hopelessly retarded" children from an orphanage to homes outside the institution and observed significant increases in IQs. A follow-up study conducted 30 years later (Skeels, 1966) showed that the children maintained these gains.

Although there is imperfect agreement about the heritability of intelligence, a great deal is known about specific environmental factors which can affect intelligence. There are, for example, a number of prenatal influences including rubella, anoxia, alcohol consumption by the pregnant mother (fetal alcohol syndrome) cigarette smoking by the mother, and maternal health status. Metabolic disorders

BEST COPY

such as galactosemia or PKU can affect development if not treated. Toxins such as lead poisoning can have an effect as can malnutrition. Birth weight is also an important factor. For example, Wiener and Milton (1970) reported an increase in mental and educational impairment of babies weighing less than 2,000 grams (5 pounds) at birth.

Family and child rearing practices can affect intellectual development also. McCall, Appelbaum and Hogarty (1973) found that children from homes that had minimal stimulation and a severe punishment regime showed decreases in IQ. Similarly, Wolf (1966) found a remarkably high correlation of .87 between IQ and the quality of the environment. Birth order and family size have proven to be significant factors. Zajonc (1976) explained that later-born children of large families frequently obtain lower IQs than early-born children. There is evidence that sex is a correlating factor. In an extensive study of sex differences, Maccoby and Jack (1974) found that girls have somewhat better verbal ability than boys and that boys are better in mathematical and spatial abilities due to the values of the culture. There is also evidence that socioeconomic level is highly associated with intelligence. Tyler (1965) stated that, "the relationship of measured intelligence to socioeconomic level is one of the best documented findings in mental test history" (p. 336). Eysenck (1973), for example, studied the obtained IQs of hundreds of people in various occupations and found that people in prestigious, high-level professions such as physicians, lawyers, and professors had a mean IQ of 130 whereas people in semiskilled and low status occupations such as laborer, farmhand or miner had a mean IQ of 90.

## INTELLIGENCE TESTS

Intelligence is, supposedly, measured by intelligence or mental ability tests. Simplistically, an IQ test is a systematic means of comparing samples of behavior with those of a norm group. Consequently, the tests are standardized,

norm-referenced instruments used to measure acquired behaviors believed to be correlated with or indicative of intelligence. More than 200 such tests have been developed (Buros, 1979), some of which are good instruments and others quite poor. There are two major types - group administered and individually administered. Most criticism has justifiably been focused on the group administered tests such as the California Test of Mental Maturity.

Although all IQ tests are different, they do share common characteristics. All such tests yield a score. These scores may be expressed variously as a MA, IQ, Cognitive Index, Mental Index etc. All IQ tests require some kind of response be it verbal, such as answering a question, or nonverbal, such as pointing to a picture or arranging blocks to form a pattern. All such tests are standardized which means that the test must be administered and scored in a prescribed manner. All tests are norm-referenced meaning that the test, during development, was given to a representative sample from which performance standards (norms) were established. All tests provide information on administration, criteria for scoring the test, and tables of norms for interpretation. Further, the tests share in common attributes of validity and reliability. Finally, all tests have inherent error called the "standard error of measurement". This statistic, based on the standard deviation and reliability of the test, provides interpretive information regarding the range of scores in which the "true IQ" probably lies.

While there are common elements among IQ tests, there are also significant differences. One major difference is the population for whom the test was designed. Most tests encompass a specific age range. For example, the Merrill-Palmer Scale was developed for use with children ranging in age from 18 months to six years while the Wechsler Intelligence Scale for Children-Revised (WISC-R) is for children ranging in age from six to seventeen years. Further, ability-specific tests have been developed such as the Perkins-Binet for blind children or the Hiskey-Nebraska for Deaf children.

The content of tests is also different. Some tests like the Leiter International Performance Scale are totally non-verbal whereas others, such as the Stanford-Binet, are highly verbal. Some (Peabody Picture Vocabulary Test) provide an IQ based only on the measurement of one skill (identifying pictures) although others, such as the Kaufman Assessment Battery have a number (13) of different subtests measuring myriad skills and abilities.

Additional differences include length of administration time. For example, the Sleason Intelligence Test takes only about 10 minutes to give whereas the Kaufman Assessment Battery can take up to one and one-half hours. Tests also vary greatly relative to quality of construction, standardization, development of representative norms, and validity. Therefore, some tests like the WISC-R are considered to be excellent instruments which provide highly reliable scores whereas others (Perkins-Binet) have little, if any, proven validity or reliability.

In explaining what IQ tests are, it is equally important to point out what they are not. Sattler (1982) provided a succinct listing:

1. IQ tests do not measure innate intelligence
2. IQs are not fixed and immutable
3. IQ tests do not provide perfectly reliable scores
4. IQ tests do not measure everything known about a person's intellectual ability.

In summary, there is acknowledged controversy over the use of IQ tests as well as the whole concept of intelligence. Some critics wonder if intelligence can even be measured although most psychologists believe that IQ tests can provide some useful information regarding cognitive functioning. Further, there is considerable variation in the quality of IQ tests which must be considered.

## FACTORS AFFECTING IQ MEASUREMENT

A number of factors have a direct effect on the measurement of intelligence. One of the most fundamental is the examiner. Most intelligence tests (except group-administered tests) are restricted meaning that they can only be given by

qualified examiners, usually psychologists who meet standards established by the American Psychological Association (APA). However, some states such as Georgia or Alabama employ psychometrists with minimal training. Further, other professionals such as teachers, speech therapists, or educational diagnosticians sometimes give IQ tests.

While most psychologists are competent in administering IQ tests to most children, a notable exception is children with sensory disabilities and other severely handicapping conditions. For example, Levine (1974) investigated psychologists' competence in testing deaf children. She learned that 90 percent of psychologists who tested these children could not communicate with the child and 65 percent had had no prior experience or training in giving tests to deaf children. Similarly, Alford (1980) found that the majority of psychologists in his nation-wide study knew very little about the appropriate assessment of a blind child. Most psychologists are not qualified to test non-English speaking children as was documented in the case of Larry P. v W. Riles (Baca & Cervantes, 1984). Further, it is not uncommon for many psychologists to refer to children with severe handicaps as "untestable". Consequently, the training and experience of the examiner have a significant impact on test results.

Examiner bias can also be a factor affecting IQ. For example, Ross and Salvia (1975) found that the physical attractiveness of the student being tested was a biasing factor; attractive children tended to receive somewhat higher scores than less attractive children. In a similar vein, examiners sometimes differ in the credit they give to a response on the test, dependent upon their perceptions of the child (Sattler, 1982). Sattler also pointed out that female examiners have been found to obtain slightly higher scores than male examiners. Some researchers have claimed that the race of the examiner is a biasing factor. However, Sattler (1982) reviewed 29 studies available on this issue and reported that in

25 of the studies, no significant relationship was found between the race of the examiner and the examinee's score.

Another factor involves inappropriate matching of the test and the person being tested. Examples include the administration of a test to a child who is younger or older than the norm group or to children who are not comparable to the test norm group (Sylvia & Ysseldike, 1978). Another common, yet inappropriate, procedure is to give only part of the test.

The age of the person being tested is a factor. Generally, the younger the child the less reliable the score. Scores obtained on children less than five years of age are not too reliable and scores obtained on infants have practically no reliability at all (McCall, 1977).

A further dimension involving IQ measurement is the interpretation of the obtained score. Unfortunately, the invention of the IQ by Terman has contributed greatly to misinterpretation. Many people believe, for example, that a child with an IQ of 109 is smarter than a child with an IQ of 92 which is an incorrect interpretation in that both children scored within the average range so are considered to be of average intelligence. IQ tests simply are not sensitive enough at present to make such quantitative comparisons. In interpreting scores, one must keep the standard error of measurement (SEM) in mind. The SEM of the WISC-R is roughly four points. A further compounding factor is that of "confidence intervals" or the probable extent to which an obtained score reflects the "true" abilities of the child. Suppose a student attained a WISC-R IQ of 98 which, in interpretation, means that there is a 95 percent probability the "true" score lies somewhere between the range of 86 to 120. The common mistake is to assume that IQ is quantitatively descriptive when, in fact, the score is only meaningful in terms of ranges of scores and confidence intervals.

Unintentional errors can affect accurate measurement. These errors include the incorrect computation of the child's age, use of the wrong table of norms,

scoring the test incorrectly, or even adding up the points earned incorrectly. Subjective factors such as the child's physical and mental condition at the time of testing, the physical environment, interruptions during testing, the quality of rapport, or the examiners style, all can affect performance (Louttit, 1957).

Poor tests are a major contributing factor in intelligence measurement. An unreliable test measures error, not ability, while an invalid test does not measure what it purports to measure. Sylvia and Ysseldyke (1978) provided a listing of commonly-used tests which do not meet APA Standards for Educational and Psychological Tests. Included in this listing is the Stanford-Binet which is one of the most widely used IQ tests for the purpose of detecting mental retardation. Scores obtained from this test since 1916 have been the primary basis for the placement of children in classes for the mentally retarded. The newly published Perkins-Binet (Davis, 1981) provides little information on the validity and reliability of the test and no information regarding standard deviation. Buros (1972), in speaking about the more than 2,500 tests reviewed in his Seventh Mental Measurements Yearbook, stated, "at least half of the tests currently on the market should never have been published. Exaggerated, false or unsubstantiated claims are the rule rather than the exception (p. xxvii).

There are, without question, valid criticisms of IQ tests and the manner in which they are sometimes used. However, the fact remains that IQ tests, properly used by qualified and experienced clinicians, can be extremely useful. One improtant use of these tests is diagnosis and classification, especially for the purpose of providing needed educational services. Whatever ones position is regarding labeling and classification, IQ tests have proven to be the most important factor in appropriate diagnosis. For example, the definition of mental retardation (Grossman, 1973) is based on IQ. Similarly, definitions of learning disabilities, "giftedness", and emotional disturbance are based, in part, on IQ. No one can deny that mistakes have been made - that children have been mislabeled on the

basis of test scores. Little attention is given, however, to the thousands of children who receive needed special education as a result of appropriate assessment.

Cottle (1974) provided a poignant example of the usefulness of IQ. He related the case of B.J. Harris, a 10 year old black boy whom he encountered in an urban school. B.J. had been placed in a class for mentally retarded children yet he did not seem to be retarded to Cottle. A check of his records revealed that the child had been arbitrarly assigned an IQ and placed in the special class on the basis of teacher recommendation. He had never been tested. A further check of records turned up seven additional children in the same circumstances. Upon administration of an IQ test, B.J. and six of the other children were found to have normal intellectual abilities so were removed from the class and returned to their fourth grade class. Weiner and Stewart (1984) related a similar example of a child who had been assigned an IQ of 82 by a speech therapist on the basis of a picture-identification test. When given an appropriate test by a qualified examiner, the girl was found to have an IQ of 126. Consequently, IQ tests can also prevent the mislabeling of children or, at minimum, help to assure proper classification.

Another reason for the use of IQ tests is that it is required by law in the identification of handicapped children for educational purposes. The Education of All Handicapped Children Act of 1975 (P.L.94-142) contains a number of requirements regarding the psychoeducational assessment of handicapped children. The most salient are listed as follows:

1. The child must be assessed in all areas relating to a disability including health, vision, hearing, social/emotional status, general intelligence, academic performance, communicative status, and motor abilities.

2. Tests shall be administered by trained and qualified personnel

3. No single procedure (test) is to be used as the sole criterion for determing a handicapping condition.

BEST COPY

A plethora of research has been reported regarding the usefulness of IQ tests. French (1974) and Martinson (1973), for example, found that the single best method available for the identification of children as "gifted" is a standardized, individually-administered IQ test. Butcher (1968) reported that IQ is an excellent predictor of scholastic achievement and others, such as Weiner and Stewart (1984), indicated that IQ is the best known long-range predictor of subsequent success and adjustment. Willerman (1979) wrote, "if one were looking for a single childhood augury of outstanding later accomplishment, one could not do better than to obtain an intelligence test measure on the subject" (p. 332). In fact, IQ has a larger collection of correlates that are predictive of success in a wide variety of circumstances than any other variable (Anderson & Messick, 1974).

Although IQ tests have proven to be highly useful, there is an undeniably sad and unfortunate history of abuse. This abuse is well illustrated by several quotes of noted psychologists who, in the earlier part of this century, testified before congress regarding immigrants entering the United States.

> "Goddard reported that, based upon his examination of the great mass of average immigrants, 83% of Jews, 80% of Hungarians, 79% of Italians, and 87% of Russians were feebleminded."
>
> "The fact that the immigrants are illiterate or unable to understand the English language is not an obstacle...Beta (an intelligence test) is entirely objective".
>
> "All mental testing upon children of Spanish-American descent has shown that the average intelligence of this group is even lower than the average intelligence of the Portuguese and Negro children, and the average intelligence of the French Canadian group in our data approaches the level of the average Negro intelligence".
>
> (Kamin, 1975, p. 319)

Today, fortunately, no one can rationally believe that 87 percent of Russians or 79 percent of Italians are mentally retarded and few psychologists would claim that an IQ test is "entirely objective". However, many critics still maintain that these tests discriminate against minority groups although the data are

equivocal. For example, Williams (1971) claimed that, "IQ and achievement tests are nothing but updated versions of the old signs down South that read, 'for Whites only'" (p. 64), and there is some evidence to support his claim. Mercer (1973), as well as numerous others, has pointed out that there is a disproportionate representation of minority students in classes for the mentally retarded. Other researchers such as Shuey (1966) have reported lower mean IQs of minority groups, in comparison with White middle class children; this gap is generally considered to be about 15 points. Further evidence of test bias was provided by the 1978 court case of Larry P. v. Riles (Payne & Patton, 1981) in which the judge restrained the use of IQ tests in the placement of black students in classes for the mentally retarded. Critics of IQ tests, citing this court case, usually fail to mention, however, a similar 1980 case (PASE v Hannon) in Illinois where the judge permitted continued use of IQ tests with black students since he found little evidence of bias in the tests.

In presenting their case of test bias, critics often point to specific test items which they believe to be discriminatory toward minority children. Williams (1971) identified one such blatantly biased item on the WISC Comprehension subtest. "What should you do if a child smaller than you begins to fight with you?" He reasoned that the question reflected middle class mores so was inappropriate for minority children. Miele (1979), however, found that the item proved to be easier for black children than for white children. Another finding which tends to discredit the notion of test cultural bias is that Japanese children obtain higher mean scores on many of the WISC Performance (non-verbal) subtests than do Anglo American children even though there are obvious cultural differences.

The most common explanation of demonstrated differences in mean scores among various cultural and minority groups is that the tests simply reflect the characteristics of the groups. Green (1978) provided a summarizing statement:

> The problem of poor achievement of black inner city children is real. It has not been caused by tests. Tests have been helpful in documenting the severity of the educational deficits. The tests are not bigoted villains but color-blind measuring instruments that have demonstrated a social problem to be solved. (p. 669)

## PAST, PRESENT, AND FUTURE TRENDS IN INTELLIGENCE MEASUREMENT

A great deal of research has been accomplished over the past few years aimed at improving tests and in developing new techniques. It is too early to determine how successful these new approaches will be although some research is available.

One approach which has been tried is the development of "cultural-fair tests". These tests attempt to reduce or eliminate those factors which are believed to be culturally loaded such as language. Instead, the tests emphasize non-verbal content. Examples include the Culture Fair Intelligence Test and The Progressive Matrices. According to Arvey (1972), however, this approach has not been successful. He noted that ethnic minorities do no better on these tests than on the more conventional tests such as the WISC-R. Sattler (1982) further pointed out that the tests have no greater validity for ethnic minorities than do more verbally loaded tests. Another reason the tests have not proven to be useful was provided by Wesman (1968) who explained that intelligence is based, in part, on the learning experiences provided by culture and environment so no test can be created which eliminates the influence of cultural learning.

Other researchers have tried the opposite approach - the development of culture-specific tests. The idea here is not to develop a test devoid of cultural influences but, instead, to develop a test for a specific culture with corresponding cultural influences. Perhaps the best known is the Black Intelligence Test of Cultural Homogeneity or the BITCH (Williams, 1971). The test consists of 100 multiple-choice questions based on items associated with black culture. Most items are, in reality, temporal slang used by inner city blacks. Several examples

are as follows:

What is a deuce and a quarter?
a. money c. house
b. car d. dice

Boot refers to a:
a. cotton farmer c. indian
b. black d. Vietnamese citizen

Williams, himself, conducted research with the BITCH and in one study reported low correlations between the test and the California Achievement Test. Other researchers, however, failed to support even his reported moderate correlations. Long and Anthony (1974), for example, compared WISC and BITCH scores of 30 black students enrolled in a class for mentally retarded students. The students' mean scores on the BITCH were at the first percentile while their mean score on the WISC was at the third percentile. Thus, these black students performed more poorly on the BITCH than on the WISC which tends to discredit the notion that the BITCH is less culturally biased than more conventional tests. Referring to the BITCH, Milgram (1974) stated that the test may be useful for developing black pride but it is useless as a predictor of success.

Another, more successful, trend has been the reduction or elimination of age, sexual, and racial bias. These attempts range from such simplistic modifications as including more pictures of females and minorities to more sophisticated standardization procedures. For example, the WISC was standardized only on white children whereas the revised WISC standardization sample included all racial, socioeconomic, and geographic groupings based on the U.S. census. Thus, the newer tests are technically better and provide a more balanced representation in the norm group.

Test content is also being changed due to greater awareness of ethnic and sexual biases. One example was given by Alley and Foster (1978) who explained that the old WISC contained the following question: "When is Washington's birth-

day?" These authors noted that the question was appropriate for white children yet there was the possibility that black children might respond to the question with the answer, "April 5" which, according to the manual, is incorrect even though that is the correct birthdate of Booker T. Washington. This question was eliminated from the revised WISC.

One of the most innovative and controversial approaches is SOMPA or the System of Multicultural Pluralistic Assessment by Mercer (1977). This is a battery of nine different measures which includes the WISC-R, other tests, and inventories for use with black and hispanic children. By means of a complicated system of weighing scores based on culture and other factors, an adjusted IQ is obtained which Mercer calls an Estimated Learning Potential (ELP). A number of critics have pointed to a lack of demonstrated validity and adequate national norms. A particularly vocal critic is Clarizio (1979) who charged that the technique tends to exclude children from special education programs when such special assistance is needed. He explained that, "Mercer's scale will prove far more discriminatory than the use of traditional IQ tests" (p. 87). However, the states of California and Louisiana have already adopted its use for determining the placement of minority children in special educational programs rather than the traditional IQ (Rice, 1979).

In addition to the development of new tests and the revisions of old ones, new techniques are being investigated to measure intelligence in different ways. One such technique is neurometrics or the measurement of brain electrical activity in response to stimuli by use of the EEG. The brain's response or "evoked potential" provides quantitative measures which can be statistically correlated with IQ tests as an indication of learning ability (Rice, 1979). Sattler (1982), however, citing the research of Rhodes, Dustman and Beck (1969) indicated that such an approach is useless for the appraisal of intelligence.

In a similar vein, researchers are using other physiological measures such as

heart and muscle responses to situations. Rice (1979) provided an example in which a child watches a toy car roll down a ramp which knocked down a styrofoam doll. The same procedure was repeated six times so the child learned what to expect. However, on the seventh trial the car did not knock the doll over which violated the child's expectation resulting in a change of heart rate which researchers believe provides a relatively unbiased measure of mental ability.

Cognitive psychologists are also investigating and developing new theories on how the brain functions which could ultimately have a significant impact on IQ tests of the future. Sternberg (1979), for example, is studying information processing or the cognitive steps involved in problem solving by the use of analogy problems similar to the following example:

> Washington is to one as Lincoln is to:
> a. five b. ten c. fifteen d. fifty

Another area of research involves cognitive style or distinctive patterns of thinking. This concept includes such dimensions as convergent versus divergent thinking; field independent versus field dependent; right brain versus left brain; and masculine versus femine cognitive differences.

A 1979 issue of Intelligence contained various experts' predictions of what IQ tests will be like in the year 2,000. One contributer (Turnbull, 1979) forsees evolutionary rather than quantum changes in IQ tests although he though the term "IQ" will have been abandoned by that time. Resnick (1979) predicted that the tests will become even more important in the future, especially for children in special education programs. She thinks, however, that the tests of the future will involve more measurement of the learning process rather than simply performance. A third predictor (Horn, 1979) took a more pragmatic view by stating that the tests used in the year 2,000 will be similar to those used today and 40 years ago. However, he predicted that cognitive psychologists will learn more about intelligence and, consequently, how to measure it.

Although no one can predict the future, past and present trends seem to suggest that more, rather than less, IQ testing will occur. Further, in view of the improved technical quality of tests today in comparison with those of years ago, it is probable that the techniques of the future will be even more precise and fairer. The history of IQ testing makes one fact very clear - that, regardless of critics, the IQ test is here to stay. Consequently, we have the choice of expending our energy in trying to unsuccessfully ban IQ tests, or we can apply our knowledge and efforts in developing better tests and in reducing or eliminating abuses. This choice, however, must be made on the basis of reason and facts; not on the basis of ignorance or histrionics.

BEST COPY

## REFERENCES


Alford, D.W. (1980) Knowledge needed and possessed by school psychologists in the psychoeducational assessment of visually impaired children as perceived by school psychologists and teachers of visually impaired children. Unpublished Doctoral Dissertation, University of Pittsburgh.

Alley, G. & Foster, C. (1978) Nondiscriminatory testing of minority and exceptional children. Focus on Exceptional Children, 9, 1-14

Anderson, S. & Messick, S. (1974) Social competency in young children. Developmental Psychology, 10, 282-293.

Arvey, R.D. (1972) Some comments on culture fair tests. Personnel Psychology, 25, 433-448.

Baca, L.M. & Cervantes, H.T. (1984) The Bilingual Special Education Interface, 8, 136-160.

Buros, O.K. (Ed.) (1972) The seventh mental measurements yearbook. Highland Park, NJ: Gryphon Press.

Buros, O.K. (Ed.) (1979) The eighth mental measurement yearbook. Highland Park, NJ: Gryphon Press.

Butcher, H.J. (1968) Human intelligence: Its nature and assessment. London: Methuen.

Clarizio, H.F. (1979) In defense of the IQ test. School Psychology Digest, 8, 79-88.

Cottle, T.J. (1974) If B.J. Harris' IQ score was faked. Why? and how many more might there be? Learning, 8, 36-40.

Cronbach, L.J. (1975) Five decades of public controversy over mental testing. American Psychologist, 30, 1-14.

Currence, C. (1984) Education Week, 1v, p.33

Davis, C.J. (1981) Perkins-Binet tests of intelligence for the blind. Watertown, MA: Perkins School for the Blind.

Eysenck, H.J. (1973) The measurement of intelligence. Baltimore: Williams and Wilkins.

French, J.L. (1974) The gifted. In M.V. Wisland (Ed.), Psychoeducational diagnosis of exceptional children. Springfield, Ill: Charles C. Thomas.

Green, B.F. (1978) In defense of measurement. American Psychologist, 33, 664-670

Grossman, H.J. (1973) Manual on terminology and classification in mental retardation. Baltimore: Garamond/Pridemark Press.

Horn, J.L. (1979) Trends in the measurement of intelligence. Intelligence, 3, 229-239.


BEST COPY

Jensen, A.R. (1969) How much can we boost IQ and scholastic achievement. Harvard Educational Review, 39, 1-123.

Kamin, L.J. (1974) The science and politics of IQ. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kamin, L.J. (1975) Social and legal consequences of IQ tests as classification instruments: Some warnings from our past. Journal of School Psychology, 13, 317-323.

Levine, E.S. (1974) Psychological tests and practices with the deaf: A survey of the state of the art. The Volta Review, 76, 298-319.

Loehlin, J.C., Lindzey, G., & Spuhler, J.N. (1975) Race differences in intelligence. San Francisco: W.H. Freeman.

Long, P.A. & Anthony, J.J. (1974) The measurement of mental retardation by a culture-specific test. Psychology in the Schools, 11, 310-312.

Louttit, L.M. (1957) Clinical psychology of exceptional children (3rd. ed.). New York: Harper.

Maccoby, E.E. & Jack, C.N. (1974) The psychology of sex differences. Stanford, CA: Stanford University Press.

Martinson, R.A. (1973) Children with superior cognitive abilities. In L.M. Dunn (Ed.), Exceptional children in the schools: Special education in transition. New York: Holt, Rinehard and Winston.

McCall, R.B. (1977) Childhood IQ's as predictors of adult educational and occupational status. Science, 197, 482-483.

McCall, R.B., Appelbaum, M.I. & Hogarty, P.S. (1973) Developmental changes in mental performance. Monographs of the Society for Research in Child Development, 38, 1-83.

Mercer, J.R. (1973) Labeling the mentally retarded. Berkeley: University of California Press.

Mercer, J.R. (1977) System of Multicultural Pluralistic Assessment: Technical manual. New York: Psychological Corporation.

Miele, F. (1979) Cultural bias in the WISC. Intelligence, 3, 149-164.

Milgram, N.A. (1974) Danger: Chauvinism, scapegoatism, and euphemism. In G.J. Williams and S. Gordon (Eds.), Clinical child psychology: Current practices and future perspectives. New York: Behavioral Publications.

Payne, J.S. & Patton, J.R. (1981) Mental retardation. Columbus, OH: Charles C. Thomas.

Resnick, L.B. (1979) Social and scientific conceptions of intelligence. International Review of Applied Psychology, 24, 131-138.

Rice, B. (1979) Brave new world of intelligence testing. Psychology Today, 13, 27-41.

Rhodes, L.E., Dustman, R.E. & Beck, E.C. (1969) The visual evoked response: A comparison of bright and dull children. In J.M. Sattler, Assessment of children's intelligence and special ability (2nd Ed.). Boston: Allyn and Bacon.

Sattler, J.E. (1982) Assessment of children's intelligence and special abilities. Boston: Allyn and Bacon.

Scarr, S. (1978) From evolution to Larry P., or what shall we do about IQ tests? Intelligence, 2, 325-342.

Shuey, A.M. (1966) Testing of negro intelligence (2nd Ed.). New York: Social Science Press.

Skeels, H.M. (1966) Adult status of children with contrasting early life experiences. Monographs of the Society for Research in Child Development, 31, (3)

Skodak, M. & Skeels, H.M. (1949) A final follow-up study of one hundred adopted children. Journal of Genetic Psychology, 75, 85-125.

Sternberg, R. (1979) Stalking the IQ quark. Psychology Today, 13, 42-54.

Sylvia, J. & Ysseldike, J.E. (1978) Assessment in special and remedial education. Boston: Houghton Mifflin.

Turnbull, W.W. (1979) Intelligence testing in the year 2000. Intelligence, 3, 275-282.

Tyler, L.E. (1965) The psychology of human differences (3rd Ed.). New York: Appleton-Century Croffs.

Vernon, P.E. (1969) Intelligence and cultural environment. London: Methuen.

Vernon, P.E. (1979) Intelligence: Heredity and environment. San Francisco: W.H. Freeman.

Wechsler, D. (1958) The measurement and appraisal of adult intelligence (4th Ed.). Baltimore: Williams and Wilkins.

Wesman, A.G. (1968) Intelligence testers. American Psychologist, 23, 267-274.

Wiener, G. & Milton, T. (1970) Demographic correlates of low birth weight. American Journal of Epidemiology, 91, 260-272.

Weiner, E.A. & Stewart, B.J. (1984) Assessing individuals: Psychological and educational tests and measurements. Boston: Little Brown.

Willerman, L. (1979) The psychology of individual and group differences. San Francisco: W.H. Freedman.

Williams, R.L. (1971) Abuses and misuses in testing black children. Counseling Psychologist,2, 62-73.

Wolf, R. (1966) The measurement of environments. In A. Anastasi (Ed.), Testing problems in perspective. Washington, D.C.: American Council on Education.

Zajonc, R.B. (1976) Family configuration and intelligence. Science, 192, 227-236.